

REC'D PCT/PTO 27 5 JUL 2004

PCT/JP03/00261

BEST AVAILABLE COPY

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

15.01.03

XX

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2002年 7月23日

出 願 番 号

Application Number:

特願2002-213695

[ST.10/C]:

[JP2002-213695]

出 願 人

Applicant(s):

鈴木 俊明

REC'D 14 MAR 2003

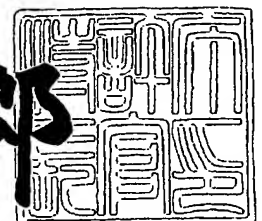
V. PO PCT

PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2003年 2月25日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

太田信一郎



出証番号 出証特2003-3010620

【書類名】 特許願

【整理番号】 P02040

【提出日】 平成14年 7月23日

【あて先】 特許庁長官殿

【国際特許分類】 A61B 5/00

【発明者】

【住所又は居所】 東京都国立市西 2 - 2 9 - 8 9

【氏名】 鎌谷 直之

【発明者】

【住所又は居所】 神奈川県鎌倉市玉縄 5 - 1 2 - 9

【氏名】 田中 順治

【発明者】

【住所又は居所】 神奈川県藤沢市川名 8 4 9 番地 1 0 号クリオ藤沢参番館  
5 0 6

【氏名】 鈴木 俊明

【特許出願人】

【識別番号】 502060360

【氏名又は名称】 鈴木 俊明

【先の出願に基づく優先権主張】

【出願番号】 特願2002- 42355

【出願日】 平成14年 1月15日

【代理人】

【識別番号】 100097113

【弁理士】

【氏名又は名称】 堀 城之

【手数料の表示】

【予納台帳番号】 044587

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0209699

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 SNP 特定方法

【特許請求の範囲】

【請求項 1】 疾患易罹患性や薬剤応答性に関する SNP 特定方法であって

SNP の解析の対象となる塩基配列領域において、予め走査領域を設定する第 1 のステップと、

前記走査領域うち目標 SNP を含む局所的な領域に段階的に絞込む第 2 のステップと、

絞込まれた前記局所的な領域から前記目的 SNP を特定する第 3 のステップとを有する

ことを特徴とする SNP 特定方法。

【請求項 2】 前記第 2 のステップは、前記目的 SNP を特定するためのマーカー SNP を確定し、前記走査領域を段階的に絞込むステップを含むことを特徴とする請求項 1 に記載の SNP 特定方法。

【請求項 3】 前記第 2 のステップは、ハプロタイプ解析などの統計解析を用いて前記マーカー SNP を確定することを特徴とする請求項 2 に記載の SNP 特定方法。

【請求項 4】 前記第 1 のステップは、機能が解明されている遺伝子又は機能の予測ができる染色体等の限定されたゲノム領域において、前記塩基配列領域での走査領域を設定するステップを含み、

前記第 2 のステップは、

前記走査領域の中から、タイピングする SNP のグループ選定を行い、wet プロセスなどによる SNP タイピングを行う第 4 のステップと、

前記 SNP タイピングによるタイピング・データに基づき、前記走査領域において前記ハプロタイプ解析の各々組合せの出現する確率を統計量として求める第 5 のステップと、

求められた前記統計量と予め設定又は測定された基準統計量とを比較し、予め設定された閾値を越える前記統計量と前記基準統計量との乖離がある場合、前記

閾値を越えて乖離した位置に該当する領域に前記マーカー S N P が含まれると判断する第 6 のステップとを含む

ことを特徴とする請求項 3 に記載の S N P 特定方法。

【請求項 5】 前記第 3 のステップは、

前記乖離が第 1 の閾値以下／未満の場合、前記第 4 のステップでの S N P のグループ選定におけるタイピングの対象となる S N P 数を所定の割合で増加させ前記第 5 のステップを繰り返す第 7 のステップと、

前記乖離が前記第 1 の閾値を越える／以上で、且つ第 2 の閾値以下／未満の場合、乖離したピークの位置を含み、前記走査領域を所定の割合で縮小した新しい走査領域を設定して前記第 5 のステップを繰り返す第 8 のステップと、

前記乖離が前記第 2 の閾値を越える／以上の場合、前記第 2 の閾値を越えて乖離した位置に該当する領域に前記マーカー S N P が含まれると判断し、乖離したピークの位置を含み、前記走査領域を所定の割合で縮小した新しい走査領域を設定して前記第 5 のステップを繰り返す第 9 のステップとを含む

ことを特徴とする請求項 4 に記載の S N P 特定方法。

【請求項 6】 前記第 9 のステップは、グループ選定される S N P が所定数以下／未満になった場合、全 DNA サンプルがタイピングされる目的 S N P を含む S N P と確定するステップを含むことを特徴とする請求項 5 に記載の S N P 特定方法。

【請求項 7】 前記第 7 のステップは、前記第 5 のステップの処理が所定の回数を越えた／以上の場合、目標 S N P が含まれないと判断して処理を中止するステップを含むことを特徴とする請求項 5 に記載の S N P 特定方法。

【請求項 8】 前記第 8 のステップは、前記第 5 のステップの処理が所定の回数を越えた／以上の場合、目標 S N P が含まれないと判断して処理を中止するステップを含むことを特徴とする請求項 5 に記載の S N P 特定方法。

【請求項 9】 前記第 2 のステップは、品質管理された工程で S N P のタイピングを行うステップを含み、前記品質管理された工程は、1 つの試料に対して、1 枚のアッセイ・プレートで 4 つの S N P タイピングを行い、カイ二乗検定などの統計手法による有意差があると認められた S N P タイピングの数が所定数を

越える場合、タイピング・データは無効と判断して前記試料のコンタミネーションなど汚染されたデータを識別することを特徴とする請求項 1 乃至 8 のいずれかに記載の SNP 特定方法。

【請求項 10】 前記第 2 のステップは、前記有意差があると認められた SNP タイピングの数が所定数の場合、有意差が認められた SNP についてのみ SNP タイピングを繰返し行い、有意差がなくなる結果が所定数続いたとき、前記タイピング／データを正しいと判定して採用することを特徴とする請求項 9 に記載の SNP 特定方法。

【請求項 11】 請求項 1 乃至 10 のいずれかに記載の SNP 特定方法における処理を実現可能なコンピュータプログラムであって、請求項 1 乃至 10 のいずれかに記載の各ステップをコード化したことを特徴とするコンピュータプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

SNP 機能解析プロセスの効率化と精度の向上、及び製薬企業などでの開発新薬の治験における SNP 機能解析手法を応用した薬剤応答性の SNP 特定方法に属する。

【0002】

【従来の技術】

従来の疾患易罹患性や薬剤応答性に関連する SNP (Single Nucleotide Polymorphism: 一塩基多型または一塩基多型となる座位) を特定する SNP 機能解析プロセスでは、コスト上の理由によって、解析したい SNP を数十から数千程度の数 (個所) に絞込んでから wet プロセス (注 1) である SNP のタイピングを行っている。

【0003】

図 7 は、従来の SNP 機能解析のプロセス・フローを示す図である。図 7 に示すように従来の SNP 機能解析においては、前段階 A (決定の対象となる開発新薬の決定)、前段階 B (解析するサンプルの収集)、段階 A (タイピング SNP

の決定)、段階B (wetプロセスによるSNPタイピング)、段階C (データの解析)、段階D: (「目標」SNPの特定) が順次行われる。

## 【0004】

これは、ヒトが持つ約300万の全SNPについてTagMan法(注2)でタイピングをした場合のコストが約2億円(サンプル一人当たりの試薬コスト)で、SNP機能の統計的な解析に必要な数百から数千のサンプルについてこの全SNPのタイピングを行えば数百億円から数千億円の非現実的なコストと大規模な解析施設等のリソースが必要となるからである。

## 【0005】

この為、通常のSNP機能解析プロセスでは、タイピングするSNP(以下、タイピングSNPと称す)を限定し、1千から1万程度のSNPに絞込んで機能の解析を行う。

## 【0006】

しかしながら、疾患易罹患性や薬剤応答性とSNPの関連の有無は、そのSNPをタイピングした結果から統計的に解明する以外に方法はない。この為、最終的に関連が解明される「目標」SNP(注3)は、予め、タイピングSNPとして1千から1万程度のSNPのグループに含まれて選定されていなければならない。これらのSNPが選定からもれた場合には、解析で関連のあるSNPは見つからず、解析プロセスを再度タイピングSNPグループの選定からやり直さなければならない。

## 【0007】

タイピングSNPを選び出す従来のやり方は、研究者が論文等の文献やゲノム関連のデータベース等を検索し、機能が既に解明しているヒト以外のゲノムと類似性したヒトの遺伝子の機能を予測するホモロジー検索等の手法を用いている。しかしながら、これらのゲノム情報には、ヒト・ゲノムの機能が完全に記載されていない。この為、このSNP機能解析プロセスの効率を決定するタイピングSNPを選び出すステップ、つまり如何に高い確率で「目標」SNPを予測できるか否かは、研究者個人の経験とスキル、そして偶然の要素に大きく依存している。

## 【0008】

また、SNP機能解析プロセスの問題のもう一つにデータの品質がある。SNP機能解析では、或る形質の発現の有無（例、作用や罹患の有無）によって区分されるサンプル・グループの間でSNPのタイピングを行い、両グループにおける各SNPのアレルの頻度を統計的に解析し、その形質を発現させる要因となるSNPを特定する。つまり、wetプロセスのタイピング・データに品質に問題があった場合、そのデータを基にしたSNP機能解析の結果は不正確なものになる。

## 【0009】

この問題は、SNPのタイピングは人間が介在するプロセスであることに起因する。SNP機能解析プロセス固有の品質の問題であるコンタミネーション、サンプルや試薬の取り違い等のオペレーション上のケアレス・ミス、これらデータ品質低下の原因となる要素の多くが人的なものであり、これもまた研究者個人の経験とスキルに大きく依存している。

## 【0010】

（注1） wetプロセスとは、SNPのタイピングを行うプロセス。現在のTaqMan法では、プレート上で血液等の遺伝子サンプルと試薬を反応・ハイブリタイゼーションさせ、その結果を光学的に測定し、最終的にサンプル個々がそのSNPで取るアレルのタイピングを行う。この工程をここではwetプロセスと称する。特定されたタイピングのデータの統計解析は、wetプロセスには含まれない。

## 【0011】

（注2） 蛍光標識したアレル特異オリゴとTaq DNAポリメラーゼによるPCR（PolymeraseChainReaction）反応とを利用したタイピング方法。

## 【0012】

（注3） 「目標」 SNPまたは「目標」となるSNPとは、疾患易罹患性や（開発新薬の）薬剤応答性の要因となるSNP、及び疾患易罹患性や薬剤応答性の指標となるSNP、以上2つのいずれかに該当するものを意味する。SNP機能



解析の目的は、これらの S N P を特定することである。

【 0 0 1 3 】

【発明が解決しようとする課題】

しかしながら、従来技術には以下に掲げる問題点があった。

【 0 0 1 4 】

タイピングの前に「目標」 S N P を予測して、これらを含んだ数百から数千程度の S N P のグループを適切且つ確実に選択することは困難であり、また、 S N P タイピング・データの品質を低下させる w e t プロセス中の人的エラーであるサンプルや試薬の取り違いやコンタミネーションの発生を防止することも極めて困難であるという問題点があった。

【 0 0 1 5 】

本発明は斯かる問題点を鑑みてなされたものであり、その目的とするところは、①マーカーとなる S N P の推定とその近傍 S N P の詳細なタイピングを繰り返すことで、「目標」 S N P が存在すると思われる塩基配列領域を段階的に絞込、最終的に「目標」 S N P を効率良く特定し、② w e t プロセスにおけるデータ品質の要因となるオペレーション上のケアレス・ミスを防止するプロセス管理手法やコンタミネーション等によって汚染されたデータを統計解析の前に排除することで、患者と非患者の統計量を比較し、 S N P 領域を絞り込む S N P 特定方法に関する技術を提供する点にある。

【 0 0 1 6 】

【課題を解決するための手段】

請求項 1 記載の本発明の要旨は、疾患易罹患性や薬剤応答性に関する S N P 特定方法であって、 S N P の解析の対象となる塩基配列領域において、予め走査領域を設定する第 1 のステップと、前記走査領域うち目標 S N P を含む局所的な領域に段階的に絞込む第 2 のステップと、絞込まれた前記局所的な領域から前記目的 S N P を特定する第 3 のステップとを有することを特徴とする S N P 特定方法に存する。

請求項 2 記載の本発明の要旨は、前記第 2 のステップは、前記目的 S N P を特定するためのマーカー S N P を確定し、前記走査領域を段階的に絞込むステップ

を含むことを特徴とする請求項 1 に記載の SNP 特定方法に存する。

請求項 3 記載の本発明の要旨は、前記第 2 のステップは、ハプロタイプ解析などの統計解析を用いて前記マーカー SNP を確定することを特徴とする請求項 2 に記載の SNP 特定方法に存する。

請求項 4 記載の本発明の要旨は、前記第 1 のステップは、機能が解明されている遺伝子又は機能の予測ができる染色体等の限定されたゲノム領域において、前記塩基配列領域での走査領域を設定するステップを含み、前記第 2 のステップは、前記走査領域の中から、タイピングする SNP のグループ選定を行い、wet プロセスなどによる SNP タイピングを行う第 4 のステップと、前記 SNP タイピングによるタイピング・データに基づき、前記走査領域において前記ハプロタイプ解析の各々組合せの出現する確率を統計量として求める第 5 のステップと、求められた前記統計量と予め設定又は測定された基準統計量とを比較し、予め設定された閾値を越える前記統計量と前記基準統計量との乖離がある場合、前記閾値を越えて乖離した位置に該当する領域に前記マーカー SNP が含まれると判断する第 6 のステップとを含むことを特徴とする請求項 3 に記載の SNP 特定方法に存する。

請求項 5 記載の本発明の要旨は、前記第 3 のステップは、前記乖離が第 1 の閾値以下／未満の場合、前記第 4 のステップでの SNP のグループ選定におけるタイピングの対象となる SNP 数を所定の割合で増加させ前記第 5 のステップを繰り返す第 7 のステップと、前記乖離が前記第 1 の閾値を越える／以上で、且つ第 2 の閾値以下／未満の場合、乖離したピークの位置を含み、前記走査領域を所定の割合で縮小した新しい走査領域を設定して前記第 5 のステップを繰り返す第 8 のステップと、前記乖離が前記第 2 の閾値を越える／以上の場合、前記第 2 の閾値を越えて乖離した位置に該当する領域に前記マーカー SNP が含まれると判断し、乖離したピークの位置を含み、前記走査領域を所定の割合で縮小した新しい走査領域を設定して前記第 5 のステップを繰り返す第 9 のステップとを含むことを特徴とする請求項 4 に記載の SNP 特定方法に存する。

請求項 6 記載の本発明の要旨は、前記第 9 のステップは、グループ選定される SNP が所定数以下／未満になった場合、全 DNA サンプルがタイピングされる

目的 SNP を含む SNP と確定するステップを含むことを特徴とする請求項 5 に記載の SNP 特定方法に存する。

請求項 7 記載の本発明の要旨は、前記第 7 のステップは、前記第 5 のステップの処理が所定の回数を越えた／以上の場合、目標 SNP が含まれないと判断して処理を中止するステップを含むことを特徴とする請求項 5 に記載の SNP 特定方法に存する。

請求項 8 記載の本発明の要旨は、前記第 8 のステップは、前記第 5 のステップの処理が所定の回数を越えた／以上の場合、目標 SNP が含まれないと判断して処理を中止するステップを含むことを特徴とする請求項 5 に記載の SNP 特定方法に存する。

請求項 9 記載の本発明の要旨は、前記第 2 のステップは、品質管理された工程で SNP のタイピングを行うステップを含み、前記品質管理された工程は、1 つの試料に対して、1 枚のアッセイ・プレートで 4 つの SNP タイピングを行い、カイ二乗検定などの統計手法による有意差があると認められた SNP タイピングの数が所定数を越える場合、タイピング・データは無効と判断して前記試料のコンタミネーションなど汚染されたデータを識別することを特徴とする請求項 1 乃至 8 のいずれかに記載の SNP 特定方法に存する。

請求項 10 記載の本発明の要旨は、前記第 2 のステップは、前記有意差があると認められた SNP タイピングの数が所定数の場合、有意差が認められた SNP についてのみ SNP タイピングを繰返し行い、有意差がなくなる結果が所定数続いたとき、前記タイピング／データを正しいと判定して採用することを特徴とする請求項 9 に記載の SNP 特定方法に存する。

請求項 11 記載の本発明の要旨は、請求項 1 乃至 10 のいずれかに記載の SNP 特定方法における処理を実現可能なコンピュータプログラムであって、請求項 1 乃至 10 のいずれかに記載の各ステップをコード化したことを特徴とするコンピュータプログラムに存する。

【 0 0 1 7 】

#### 【発明の実施の形態】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。

## 【0018】

## (実施の形態1)

図1は、本実施の形態1に係るSNP特定方法の概要を示すプロセス・フローである。図1に示すように、本実施の形態1に係るSNP特定方法は、決定の対象となる開発新薬の決定(ステップS1)と解析するサンプルの収集(ステップS2)と「走査領域(塩基配列領域)」の決定(ステップS3)と「タイピング」SNPの決定(ステップS4)とwetプロセスによるSNPタイピング(ステップS5)とタイピング・データによるハプロタイプの解析(ステップS6)と「マーカー」SNPの推定(解析データの決定)(ステップS7)と「目標」SNP(目標SNP)の特定(ステップS8)とを有し、ステップS3～S7を1つのサイクル(処理サイクル)として繰り返す。

## 【0019】

疾患易罹患者性または薬剤応答性に関連するSNP特定を目的として、このプロセス・モデルでは、以下8つの工程(ステップ)を実施することによって、SNPのタイピングを行う「走査領域」から段階的に絞込んで、最終的に開発新薬の薬剤応答性の有無と関連する「目標」SNPを特定する。

## 【0020】

8つのステップは以下のとおりである。まず、前段階1：決定の対象となる開発新薬の決定を行う(ステップS1)。次に、前段階2：解析するサンプルの収集を行う(ステップS2)。段階1：「走査領域」の決定を行う(予め走査領域の設定する)(ステップS3)。段階2：「タイピング」SNP(タイピングSNP)の決定を行う(ステップS4)。段階3：wetプロセスによるSNPタイピングを行う(ステップS5)。段階4：タイピング・データによるハプロタイプの解析を行う(ステップS6)。段階5：「マーカー」SNPの推定を行う(解析データの決定)(ステップS7)。段階6：「目標」SNPの特定を行う(ステップS8)。この8つのステップにおいて、段階1から段階5(ステップS3～S7)までを1つのサイクル(処理サイクル)としてこれを繰り返す。

## 【0021】

次に、図1を参照して、各ステップにおける処理を詳しく説明する。

## 【0022】

(イ) 前段階1 (決定の対象となる開発新薬の決定) : 製薬企業等によって開発された新薬のうち、開発新薬の薬剤応答性 (=作用または副作用の有無) を SNP によって検定したものを選択する。また、このプロセス・モデルでは、例えば疾患易罹患性等を対象として、SNP との関連を調べることもできる。

## 【0023】

(ロ) 前段階2 (解析するサンプルの収集) : SNP 機能解析では、或る形質の発現の有無によって区分されるサンプル・グループの間で、SNP データを比較し、その形質を発現させる要因となる SNP を特定する (ステップ S2)。例えば、糖尿病等について疾患易罹患性と SNP とのを調べる場合には、糖尿病患者のグループと、Control されたグループの間で、各 SNP のアレル頻度を統計的に分類する。この時、Control グループとして、糖尿病に罹患していないグループ、もしくは (糖尿病罹患の有無にかかわらず) 無作為に抽出された平均的なサンプル・グループ、この2つのいずれかを用いる。

## 【0024】

開発新薬の薬剤応答性に関連する SNP を特定する場合、作用もしくは副作用のあったグループと、これがなかったグループについて、段階1以降の解析を行う。

## 【0025】

無作為に抽出された平均的なサンプル・グループが存在する場合、もしくはこれらに該当する外部の SNP データが利用できる場合には、前述の2グループと合わせた3つのグループの間で、データの比較と解析を行うと、更に有効な解析を行うことが可能となる。

## 【0026】

(ハ) 段階1 (「走査領域 (塩基配列領域)」の決定) : このプロセス・モデルでは、この段階1から後に説明する段階5までを1つのサイクルとしてこれを繰り返すことによって、初期の大まかな「走査領域」からより局所的な「走査領域」へと段階的に絞込を行う。最後のサイクルでは、「走査領域」中の存在する全 SNP のタイピング・データを解析することで、最終的に「目標」SNP を特

定する。まず、走査領域の決定を行う（ステップS3）。この「走査領域」とは、「目標」SNPの存在を調べる（走査する）領域であり、ヒト・ゲノムの塩基配列上の連続した領域である。この領域の物理的な長さは、段階的に絞込まれるために可変である。

#### 【0027】

図2は、図1のステップS3、S4及びS6を説明する図である。図1及び図2を参照して、処理サイクルにおけるステップを更に、詳しく説明する。

#### 【0028】

図2の（a）は、図1のステップS3においてSNP10Dを含むゲノム領域の一例を示す。最初の「走査領域」は、ステップS3において、遺伝子やそれより大きな染色体等の大まかなレベルで規定（設定）される。これは、現状でも染色体等レベルの大まかな機能が解明されているからである。また、複数の染色体が原因となり、どの染色体が怪しい（目的SNPを含むか）かわからない場合、ある特定の染色体を除いた残りの全ての染色体（結果に男女の差が無い場合、性染色体は関係が無いので「走査領域」から除外するなどの措置をして）を対象とする場合、更に、極端な例として、全く目的SNPの絞込みに関する情報がない場合など、全染色体を取りあえず「走査領域」として解析の対象とする方法も含む。また、これより詳細な、例えば遺伝子レベルで初期の「走査領域」を設定することもできる。即ち、予め機能が解明されている染色体レベルに基づき、走査領域（1次走査領域、初期走査領域）を設定する。

#### 【0029】

2回目以降（ $(n+1)$ 回目、 $n$ は1以上の自然数とする）のサイクルにおいて、1回目（ $n$ 回目）のサイクルにおける段階5（ステップS7）から再度この段階1（ステップS3）へ戻るので、 $n$ 回目の処理サイクル（ $n$ 次処理サイクル）における段階5で連鎖不平衡が確認された領域を、 $(n+1)$ 回目の処理サイクル（ $(n+1)$ 次処理サイクル）において新しい「走査領域（ $(n+1)$ 次走査領域）」として設定する。この時、絞込まれた新しい「走査領域（ $(n+1)$ 次走査領域）」は、前回（ $n$ 回目）のサイクルにおける「走査領域  $n$ 次走査領域」の数分の1から数十分の1の長さとなる。なお、次のサイクルで、具体的にど

の程度絞込むかについては、後に述べる SNP 間の連鎖不平衡の強度等によって異なる。

#### 【0030】

(二) 段階 2 (タイピング SNP の決定) : 段階 1 で設定された「走査領域」の中から、「タイピング」する SNP のグループを選定する。

#### 【0031】

図 2 の (b) は、図 2 の (a) から選定された SNP のグループの一例を示す。これらグループに含まれる SNP は (遺伝子部位の機能やエクソン・イントロン等の区分に留意せず) 任意に選ぶことができるが、隣接する SNP 座位の間隔が出来る限り等間隔となるように選択を行う。これは、この一連の解析で SNP 座位間の連鎖不平衡を間接的に観測するため、連鎖不平衡に大きな影響を及ぼす SNP 座位間の物理的距離の差による誤差を排除するためである。

#### 【0032】

解析できる SNP 間の連鎖不平衡は、SNP 座位間の物理的距離が 1 ～ 10 万塩基程度と考えられる。この為、最初の「走査領域」が含まれる SNP が 10 万程度の染色体の全長である場合ならば、初回のタイピングは 1000 SNP 程度を行うことが望ましい。

#### 【0033】

2 回目以降の処理サイクルでは「走査領域」は段階的に絞込まれ、「走査領域」の物理的距離は短くなり、その範囲に含まれる SNP 数は減少する。タイピング SNP は、この「走査領域」の中から、数十から最大でも数百までの範囲で選択する。

#### 【0034】

図 3 は、図 1 のステップ S 5 に関するサンプル・試薬チューブ、各種プレート及び関連したデータ・スキマの一例を示す図である。

#### 【0035】

(ホ) 段階 3 (wet プロセスによる SNP タイピング) : 図 1 の wet プロセスによる SNP タイピングを行う (ステップ S 5)。即ち、選択された SNP グループについて、各サンプルの SNP タイピングを TaqMan PCR 法等

によって行う。このタイピング工程におけるコンタミネーションやサンプルの取り扱いによるデータの誤差を防ぐために、1) バーコードを用いたサンプル・試薬チューブ及びアッセイ用プレート等の世代管理、2) 「ハーディ、ワインバーグ平衡」を用いたタイピング・データの検定、これら1)、2)を行うことでタイピング・データの品質保証を行う。

## 【0036】

1) バーコードを用いたサンプル・試薬チューブ及びアッセイ用プレート等の世代管理：タイピングの工程における最も基本的で多い間違いは、サンプルや試薬を間違えて取り扱うことである。現行のSNP解析のプロセスでは、最終的にアッセイを行うタイピング機器に使用するアッセイ・プレート10APを生成するまでに、幾つかの中間プレートを生成する。その為、これらのプレートについても正しいプレートから生成されているかが管理されることが重要である。

## 【0037】

この為、サンプル及び試薬等のチューブ、これらから分注されて生成されるプレートについてバーコードを用いてID管理を行うと同時に、これらチューブ・プレート間又はプレート・プレート間の世代管理=どちらから(親)どちらが(子)生成されたかの関連を管理する。

## 【0038】

サンプル：下部に二次元バーコードを刻印されたサンプル・チューブ、及びこのサンプル・チューブが96本まで格納できるサンプル・ラックによって管理される。サンプル・ラック内の各サンプル・チューブの配置は、サンプル・ラックの底面からラック内の各サンプル・チューブのバーコードをスキャナーで読み取り、この配置をデータとして管理する。又、サンプル・ラック自身にもバーコードが添付される。

## 【0039】

試薬：下部に二次元バーコードを刻印された試薬チューブ、及びこのサンプル・チューブが96本まで格納できる試薬ラックによって管理される。サンプル・チューブと同様に、試薬ラック内の各試薬チューブの配置はバーコードをスキャナーで読み取り、この配置をデータとして管理され、試薬ラック自身にもバーコ



ードが添付される。

#### 【0040】

プレート：プレートには、マスター・プレート、試薬プレート及びアッセイ・プレート10APの3種類がある。マスター・プレートは1枚につき96のウェル（サンプルや試薬を分注するプレート上の窪み）を持ち、同じく最大96のサンプル・チューブが収納可能なサンプル・ラックからサンプルが分注される。この時、1つのサンプルは1個所のウェル（ラック内のサンプル・チューブのレイアウトと同じポジションのマスター・プレートのウェル）に分注される。この為、マスター・プレートはサンプル・ラックの「子」のデータとして管理される。試薬プレートも（マスター・プレートと同様に）96のウェルを持つが、1枚の試薬プレートには1種類の試薬のみを使うので、試薬プレートは試薬チューブの「子」のデータとして管理される。試薬チューブの試薬ラック内の配置情報を保持する理由は、分注に使う自動機器が、ラックに格納した試薬チューブの使用を要求するからである。アッセイ・プレート10APは384のウェルを持ち、最大96のサンプルに対して4 SNPのタイピングを同時に行うことができる。この1枚のアッセイ・プレート10APは、データ上4つの仮想的なバーチャル・プレートとして管理される。1つのバーチャル・プレートは最大96のサンプルに対して1つのSNPのタイピングを行うプレートで、これが仮想的に4枚合わさり1つのアッセイ・プレート10APを構成する。

#### 【0041】

2) 「ハーディ、ワインバーグ平衡」を用いたタイピング・データの検定：SNPでは、その座位の塩基が2つのバリエーション、例えばA（アデニン）またはG（グアニン）の様な2つのアレル（対立遺伝子）を持つ。全ての染色体は対になっているので、SNP座位も対になっている染色体の各々に一つずつの合計2箇所が存在する。この為、TaqManアッセイ等で観測されるSNPのパターンは、1) 一方アレル（この場合、A）のホモ接合体であるA-Aと、2) これとは異なるアレル（G）のホモ接合体G-G、3) 対立するアレルが一つずつ組合わさったヘテロ接合体A-G、以上の3つパターンになる。一方の遺伝子上でそのSNPがAを持つ確率を $\alpha$ とすれば、A-Aホモ接合体の確率は $\alpha^2$ であ

り、G-Gホモ接合体の確率は $(1-\alpha)^2$ 、A-Gヘテロ接合体の確率は $2\alpha(1-\alpha)$ となり、この関係を「ハーディ、ワインバーグ平衡」と呼ぶ。「ハーディ、ワインバーグ平衡」が成立する条件は、多くの世代の交配を経て「平衡」した状態にあるサンプル集団から無作為に、同時に統計的に「平均」に抽出されなければならない。

## 【0042】

ここで、そのサンプルのデータが「ハーディ、ワインバーグ平衡」から大きく離れた値を取った場合には、①そのデータを生成したアッセイ自体がコンタミネーション等によって「汚染」されたデータである。②タイピングされたサンプル集団が統計的には「無作為」に抽出されていない。①、②のうちのいずれかに相当すると考えられる。

## 【0043】

SNP機能解析の目的は、或る形質の発現（疾病等）の有無によって区分されるサンプル・グループの間でSNPデータを比較して、その形質を発現させる要因となるSNPを特定することである。つまり、要因となるSNPをタイピングした場合、そのグループを特徴付ける形質と因果関係のあるSNPを「期待される」以上に持っているため、「ハーディ、ワインバーグ平衡」から離れたアレルの分布を持つはずである。言葉を代えれば、この2)「ハーディ、ワインバーグ平衡」を用いたタイピング・データの検定において、該当するSNPを見つけることこそ、SNP機能解析及び本プロセス・モデルの目的である。

## 【0044】

「目標」SNPを特定するために、「ハーディ、ワインバーグ平衡」から逸脱したアレル分布のSNPから、①の「汚染」されたデータを持つSNPを除外する。

## 【0045】

このコンタミネーション等の「汚染」されたデータを識別する方法が、1枚のアッセイ・プレート10APで4つのSNPをタイピングする方法である。

## 【0046】

図4は、図3のアッセイ・プレート10AP上におけるSNPの配置パターン

の一例を示す図である。384-ウェル(384のウェル)のアッセイ・プレート10APでタイピングされるSNPのレイアウトは図4のようになる。

【0047】

コンタミネーションやPCR法における不調等のアッセイ自体の失敗が起きた場合、同じプレート上の他のSNPについても異常が発生すると考えられる。

【0048】

確率的には、要因となるSNPを偶然、複数選び出すことは極めて稀と考えられるので、同一プレート上で複数のSNPデータが「ハーディ、ワインバーグ平衡」から逸脱したケースは、タイピングの失敗と判断しそのプレートで行った全データの破棄(また、アッセイのやり直し)を行う。

【0049】

更に、タイピングしたSNP同士が連鎖不平衡する可能性を回避するため、4つのSNPを可能な限り離す＝「走査領域」を4つの区間10PT(4つのウェルのブロック)に分け、その各々の区間からタイピングするSNPを1つ選択する。

【0050】

図2の(c)は、図1のステップS6における走査領域の一例を示す。ウィンドウ10wを処理中のサイクルにおける「走査領域」の先頭から終端まで移動させて、そのウィンドウ10w内に含まれているSNPデータを解析するイメージを表している。

【0051】

(へ) 段階4(タイピング・データによるハプロタイプの解析、ステップS6)：この工程では、ハプロタイプと連鎖不平衡の2つのコンセプトを用いて、「目標」SNPの近傍のSNP(最終的に「目標」SNP自身)を特定する。

【0052】

一つの配偶子(対となる染色体の一方)上にある対立遺伝子(SNPのアレル)の組合せであるハプロタイプは、TagMan法等のSNPタイピング・アッセイで得られたデータを基に、確率的に予測される。

【0053】

このハプロタイプを確率的に予測することについて、AまたはGを取るSNP # 1、TまたはGを取るSNP # 2、TまたはCを取るSNP # 3、以上の3つのSNPが取るハプロタイプのケースから考える。

## 【0054】

例えば、サンプルXで、SNP # 1がA-Aのホモ接合体、SNP # 2がT-Gのヘテロ接合体、SNP # 3がT-Cのヘテロ接合体を取る場合には、以下の2つのケースで4通りのハプロタイプの存在が予測される。

## 【0055】

ケース1)	SNP # 1	SNP # 2	SNP # 3	確率
染色体1	A	T	T	25%
染色体2	A	G	C	25%
ケース2)	SNP # 1	SNP # 2	SNP # 3	
染色体1	A	T	C	25%
染色体2	A	G	T	25%

なお、この例の各ハプロタイプの確率は各々25%となる。

## 【0056】

更にもう一つのサンプルYが、SNP # 1がA-Aのホモ接合体、SNP # 2がT-Gのヘテロ接合体、SNP # 3がT-Tのホモ接合体を取る場合には、各々50%の確率を持つ2通りのハプロタイプが予測される。

ケース1)	SNP # 1	SNP # 2	SNP # 3	確率
染色体1	A	T	T	50%
染色体2	A	G	T	50%

## 【0057】

そして、この2つのサンプルから予測されるハプロタイプは以下の通りとなる。

SNP # 1	SNP # 2	SNP # 3	確率
A	T	T	$25\% \div 2 + 50\% \div 2 = 37.5\%$
A	T	C	$25\% \div 2 = 12.5\%$
A	G	T	$25\% \div 2 + 50\% \div 2 = 37.5\%$

A                      G                      C     $25\% \div 2 = 12.5\%$

【0058】

実際の解析は、数個から数十程度の範囲で定められた所定のSNP数を含む連続した領域をウィンドウ10wと定義し、そのウィンドウ10w内のSNPのタイピング・データ（全サンプル）から、ハプロタイプの組合せとその各々の出現する確率を統計的に求める。このウィンドウ10wに含まれるSNPの数が多すぎるとハプロタイプ個々の確率が低くなり、連鎖もしくは連鎖不平衡の有無の確認が微妙になるので、10程度のSNPを含むウィンドウ10wの定義が有効である。

【0059】

図5は、図1のステップS7（段階5）の解析データの一例を示す図である。図5では、解析データにおける検定の前半部分にあたるハプロタイプ解析から統計量に変化が現れた領域を識別する手順を示す。

【0060】

（ト）段階5（「マーカー」SNPの推定、ステップS7）：この工程では、図5に示すように、特定のハプロタイプの確率が突出している領域を見つけ出すことによって、「目標」SNPの近傍領域の推定を行う。即ち、①WINDOW（ウィンドウ）10wを移動させながらWINDOW10w内のハプロタイプを解析する。②WINDOW10wの位置による「ハプロタイプ」数等の統計量の変化をプロットする。③比較するサンプル・グループ間で、統計量に顕著な差が認められる部分を抽出する。

【0061】

段階4で解析されたハプロタイプのデータから、解析されたSNPグループ間で連鎖不平衡が見られたか否かの判定が可能となる。

【0062】

これらのSNPの間に連鎖不平衡が見られない場合、各SNPのアレルの出現頻度は「平均的」な値に落ち着き、加えて、各々のSNPが「独立」しているため、そこから統計的に求められるハプロタイプについても、特定のハプロタイプに集約されない、広く薄く分散したものになると考えられる。

## 【 0 0 6 3 】

これに対し、解析された S N P グループ間で連鎖不平衡が見られた場合、そこにはサンプル・グループを統計的に特徴付ける S N P が含まれていることであり、それらの S N P では特定のアレルの出現頻度が増大する。そして、これら S N P データを統計的に解析した結果であるハプロタイプの確率分布も、（「目標」 S N P をアッセイしなかった場合より）特定のハプロタイプに集中することが予測される。

## 【 0 0 6 4 】

この特定ハプロタイプへの集中を識別する方法として、個々のハプロタイプの出現頻度を比較するほかに、その解析データから予測されるハプロタイプの総数、これらハプロタイプの標準偏差、確率上位のハプロタイプ・グループのハプロタイプ全体に対する出現頻度の割合を「統計量」として観測し、作用・副作用の有無という形質の発現によって区分されるサンプル・グループ間でこれらと比較する。

## 【 0 0 6 5 】

しかしながら、上述した様に、「目標」 S N P を選び出して直接タイピングすることはなかなか困難である。この問題の解が、「目標」 S N P が近傍の S N P と連鎖不平衡することを応用し、「目標」 S N P の近傍領域を推定することである。連鎖不平衡した近傍の S N P は、「目標」 S N P を直接解析した場合と比べれば程度こそ弱い、やはり同じようにハプロタイプの確率分布が変化することが期待される。このような近傍の S N P は、「目標」 S N P の「マーカー」 S N P であると考えられる。即ち、対象となるサンプル（効果のあったグループ： C a s e グループ）の統計量と基準となるサンプル（効果のなかったグループ： C o n t r o l グループ）の基準統計量とを比較し、この差異が予め設定された閾値を越えた場合、該当するタイピング領域に変化があったと判断して（マーカー S N P と推定して）、該当するタイピング領域に対する所定の割合（例えば、5 分の 1 ～ 1 0 分の 1）を新しい走査領域に設定して次の処理サイクルを行う。

## 【 0 0 6 6 】

図 6 は図 1 のステップ S 7 における解析データの一例を示す図である。図 6 で

は、段階5の解析データの検定の後半部分にあたる新しい「走査領域」を定める手順を示す。

【0067】

直接「目標」SNPをタイピングすること無しに、「マーカ―」SNPのタイピング・データから、特定ハプロタイプの突出を示す領域を見つけ出し、その領域を次の「走査領域」の決定を行う段階1に戻り、再度工程を繰り返す。

【0068】

又、そのサイクルで、「走査領域」中の全てのSNPがタイピングされているのであれば、次の段階6へ進んで、その特定ハプロタイプに収束しているSNP座位が「目標」SNPとして決定される。

【0069】

この段階の目的は「マーカ―」SNPの推定であり、「走査領域」を新しく推定された「マーカ―」SNPの近傍に絞込むサイクルを繰り返すことで、「マーカ―」SNPを「目標」SNPを接近させることができる言える。なお、最終のサイクルの以前に「目標」SNPが解析から導き出されていたとしても、その時点ではまだそのSNPを「目標」SNPとして確定できてはいないので、「マーカ―」SNPとなる。

【0070】

(チ) 段階6 (「目標」SNPの特定、ステップS8) : この工程の目的は、全工程で選出された「目標」SNPと開発新薬の薬剤応答性との相関であり、その度合いを定量的に導き出すことである。

【0071】

この工程では、「目標」SNPとして確定されたSNPのアレル頻度について、サンプル・グループでの最終的な検定を行う。(カイ2乗検定や最尤法等による)

更に、サンプル・グループ以外のSNPデータ、特に東大医科研が所有する日本人の標準的なSNPに関するデータとの比較を行うことも効果的である。又、一部のゲノム研究企業で、SNPに関連するデータベースを販売する動きがあるので、これらの利用も可能である。

## 【0072】

カイ2乗検定や相関解析を用いて、「～パーセントの確率で期待される効果が得られる」、または「～パーセントの確率で重篤な副作用が発生する」等の評価検定を行う。

## 【0073】

ここで、データの汚染を識別する方法を詳しく説明する。

1) タイピング結果によって決定されたプレート中のSNPアレルの分布(1つのプレートでは最大96(実際には～92)のに対して4種類のSNPのタイピングを行うが、その各々のSNPについてのアレルの分布を意味する)

ここで「アレルの分布」とは、AA/AB/BBまたはAA/Aa/a a等の3値の何れかで表現されるSNPのタイピング・データを、そのグループ(同一プレートで同一SNPについてタイピングを行ったサンプルのSNPタイピング・データ)において3値の何れを取るサンプルの合計数を示す。(例えば、AA:23、AB:54、BB:15)

2) ハーディ・ワインバーグ平衡に基づいた理想的なSNPアレル分布:「ハーディ・ワインバーグ平衡」とは、「AA」の頻度を $\alpha^2$ とする場合、「BB」の頻度は $(1-\alpha)^2$ 、「AB」の頻度は $2\alpha(1-\alpha)$ となる状態を言う。

(この時、『「AA」の頻度』>『「BB」の頻度』)

(「A」単独の確率(=頻度)が $\alpha$ なので、両方とも「A」となる「AA」の頻度は $\alpha^2$ になり、「B」の確率=「A」じゃない確率は、 $1-\alpha$ 、「BB」の確率は $(1-\alpha)^2$ 、「AB」は「A」・「B」と「B」・「A」の2つのケースの合計の確率で $2\alpha(1-\alpha)$ になる。)

「ハーディ・ワインバーグ平衡に基づいた理想的なSNPアレル分布」とは、「AA」の頻度を $\alpha^2$ として、これから残りの「BB」の頻度と「AB」の頻度を、各々 $(1-\alpha)^2$ 及び $2\alpha(1-\alpha)$ と計算したもの。即ち、1)の例では、「AA」の頻度は $23/96=25\%$ で、「A」単独の確率は50%になる。この為、「BB」の頻度も25%でアレル分布は $96 \times 25\%=23$ 、「AB」の頻度は50%で分布は $96 \times 50\%=46$ が、「ハーディ・ワインバーグ平衡に基づいた理想的なSNPアレル分布」となる。



3) 上記2つのアレル分布の比較: 1) と 2) の2つのSNPアレルが、カイ二乗検定などで「統計的に有意差」が有るか否かを確認する。

先の例では、AA AB BBが、

1) 2 3 5 4 1 5

2) 2 3 4 6 2 3

となり、ここでは、カイ二乗検定の結果は12.41%となり、「統計的に有意差」と認められる値(5%以下の確率)より大きい為、有意差はないものと判定される。

このようにデータの汚染が識別された後、

4) 比較した結果に有意差が認められた場合、そのプレート上で同時にタイピングされたSNPの2つ以上で3)の結果に「統計的に有意差」が認められた場合、そのプレート全体のタイピングは失敗であると判定し、そのタイピング・データを廃棄する。

5) 同一プレート上におけるSNPの内1つのSNPについてのみ「統計的に有意差」が認められた場合、その有意差が認められたSNPについてのみタイピングは再度実施される(同じ解析結果が得られた場合には、そのタイピング・データは「正しい」ものとして格納される。再実行されたタイピング結果に「有意差」がなくなれば、3回目のタイピングを行う。以下この手順を繰り返し、連続して同じ結果が得られるまでタイピングを続ける。)

#### 【0074】

本実施の形態1に係るSNP特定方法は上記の如く構成されているので、以下に掲げる効果を奏する。

タイピングされたSNPの中から疾患易罹患性または薬剤応答性に関連するSNPを特定する際に、解析の対象となる塩基配列領域を大まかな領域からより局所的な領域へと段階的に絞込を行うことで、加えて、品質管理された工程でSNPのタイピングを行うことによって、最終的にこれら関連するSNPを特定することができる。

#### 【0075】

(実施の形態2)

次に、本実施の形態2に係るSNP特定方法を説明する。

【0076】

段階1（図1におけるステップS3）の「走査領域」の決定では、初期の「走査領域」を、大まかな染色体レベルより詳細な遺伝子レベルで設定することも可能である。これによって、狭い初期「走査領域」からタイピングが開始できるので、より効率的な「走査領域」が可能となる。

【0077】

段階2（図1におけるステップS4）の「タイピング」SNPの決定では、（特に、早い段階の「走査領域」において）観測するSNP座位の間隔が大きすぎた場合、連鎖不平衡はほとんど観測することができない。この連鎖不平衡が観測可能なレベルの物理的距離より短い間隔で「タイピング」SNPを選ぶことが必要である。

【0078】

初期の「走査領域」では、この間隔を、100.0SNPについて1つのSNPを選ぶ。即ち、10万塩基について1つのSNPを選ぶように初期値が設定される。しかし、これはあくまで計算値であり、その近傍に見られる連鎖不平衡の度合いによっては、より小さい値（数十SNPについて1つのSNPを選ぶ）を取るケースもある。

【0079】

段階5（図1におけるステップS7）の「マーカー」SNPの推定において、ハプロタイプの総数や出現頻度等の「統計量」の変化を比較する以外の方法として、確率上位のハプロタイプが共通SNPパターンを求めることで連鎖もしくは連鎖不平衡の有無を確定することができる。

【0080】

実施の形態1において説明したハプロタイプの確率の例では、サンプルXとサンプルYのハプロタイプの確率の和は以下の通りであった。

【0081】

SNP # 1	SNP # 2	SNP # 3	確率
A	T	T	$25\% \div 2 + 50\% \div 2 = 37.5\%$

A	G	T	$25\% \div 2 + 50\% \div 2 = 37.5\%$
A	T	C	$25\% \div 2 = 12.5\%$
A	G	C	$25\% \div 2 = 12.5\%$

この例では、このハプロタイプのグループは、75%の確率でSNP # 1がAのアレル、SNP # 3がTのアレルを取ることを示し、SNP # 1とSNP # 3に連鎖があったものと推測される。

## 【0082】

また、より複雑な下記の例では、

SNP # 1	SNP # 2	SNP # 3	SNP # 4	SNP # 5	確率
A	T	T	A	T	30%
A	T	T	A	G	25%
C	T	T	A	T	10%
C	T	T	A	G	10%
A	T	T	C	T	5%
A	T	T	C	G	5%
A	G	C	A	T	3%

## 【0083】

55%の確率で、SNP # 1がアレルA、SNP # 2がアレルT、SNP # 3がアレルT、SNP # 4がアレルAを持つハプロタイプのパターンが観測される。更に、75%の確率でSNP # 2がアレルT、SNP # 3がアレルT、SNP # 4がアレルAを持つハプロタイプのパターンが、85%の確率でSNP # 2がアレルT、SNP # 3がアレルTを持つハプロタイプのパターンが観測される。これによって、このサンプル・グループのデータからはSNP # 2とSNP # 3を中心とした連鎖が見られるものと判断される。

## 【0084】

ここで、連鎖の有無を分ける出現確率の閾値は、ウィンドウ10wで走査するSNPの数に関連する。SNP数が増えるとハプロタイプのバリエーションが増加して個々のハプロタイプが観測される確率が減るので、その結果、閾値の値も低くなる。10のSNPを観測するウィンドウ10wを使った場合、閾値として

は70%が適当と考えられる。なお、ウィンドウ10wで走査される領域（ウィンドウサイズ）は対象により変動するが、3-25 SNP程度が一般的に妥当と考えられる。

【0085】

本実施の形態2に係るSNP特定方法は上記の如く構成されているので、実施の形態1の奏する効果の他に以下に掲げる効果を奏する。

【0086】

比較する2つのサンプル・グループ間に存在する異なった連鎖、もしくは一方のサンプル・グループにしか存在しない連鎖を確認することで、「目標」SNPまたはその近傍の領域を特定できる。

【0087】

なお、本実施の形態においては、ハプロタイプ解析を用いた説明を行ったが、一般に知られている統計解析が利用できることは明らかである。また、本発明はそれに限定されず、本発明を適用する上で好適なSNP特定方法に適用することができる。また、本実施の形態を実現するため染色体レベル及びDNAレベルのSNPタイピング処理装置と、統計分析処理を行うコンピュータとを含み、一連処理ができるSNP特定システムも構成することができる。また、上記構成部材の数、位置、形状等は上記実施の形態に限定されず、本発明を実施する上で好適な数、位置、形状等にすることができる。なお、各図において、同一構成要素には同一符号を付している。

【0088】

【発明の効果】

本発明は以上のように構成されているので、以下に掲げる効果を奏する。

【0089】

以上説明したように、この発明によれば、マーカーとなるSNPを推定することで解析の対象となる塩基配列領域を大まかな領域からより局所的な領域へと段階的に絞込を行い（患者と非患者の統計量を比較し、SNP領域を絞り込む）、更に、品質管理された工程でSNPのタイピングを行うことによって、最終的に、疾患易罹患性や薬剤応答性等に関連するSNPを特定することができる。

【図面の簡単な説明】

【図 1】

本発明の実施の形態 1 に係る SNP 特定方法の概要を示すプロセス・フローである。

【図 2】

図 1 におけるステップ 3、4 及び 6 の詳細を示す図である。

【図 3】

図 1 のステップ S 5 に関するサンプル・試薬チューブ、各種プレート及び関連したデータ・スキマの一例を示す図である。

【図 4】

図 3 のアッセイ・プレート 10AP 上における SNP の配置パターンの一例を示す図である。

【図 5】

図 1 のステップ S 7（段階 5）の解析データの一例を示す図である。

【図 6】

図 1 のステップ S 7 における解析データの一例を示す図である。

【図 7】

従来の SNP 機能解析のプロセス・フローを示す図である。

【符号の説明】

10AP アッセイ・プレート

10D SNP

10PT 4つの区間

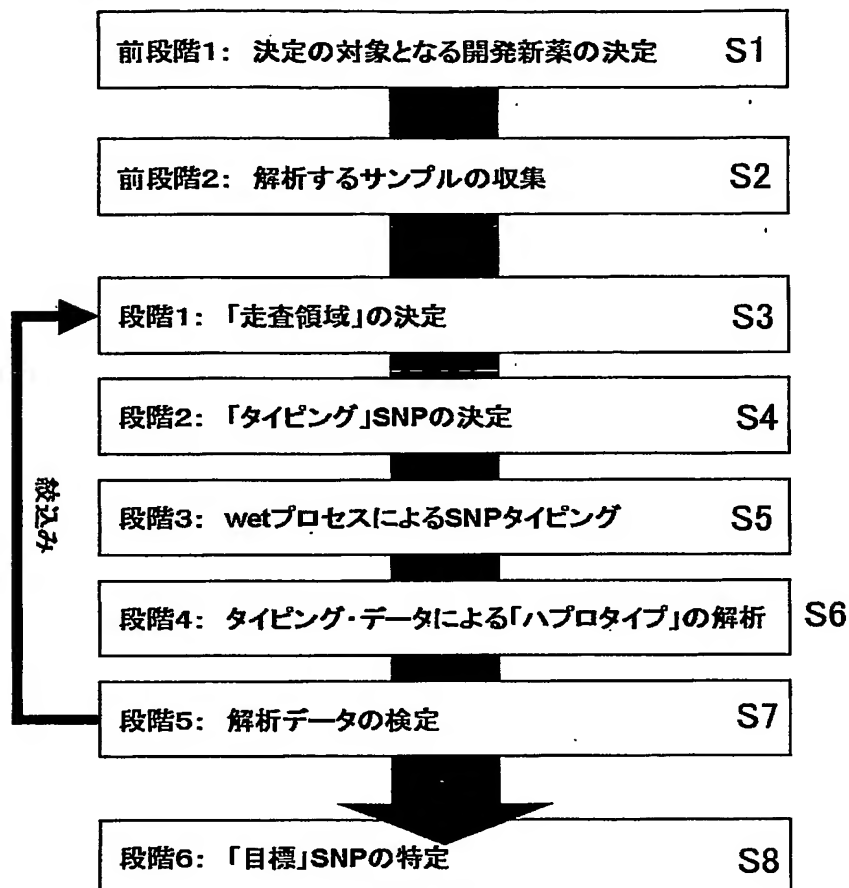
10w ウィンドウ (WINDOW)

【書類名】

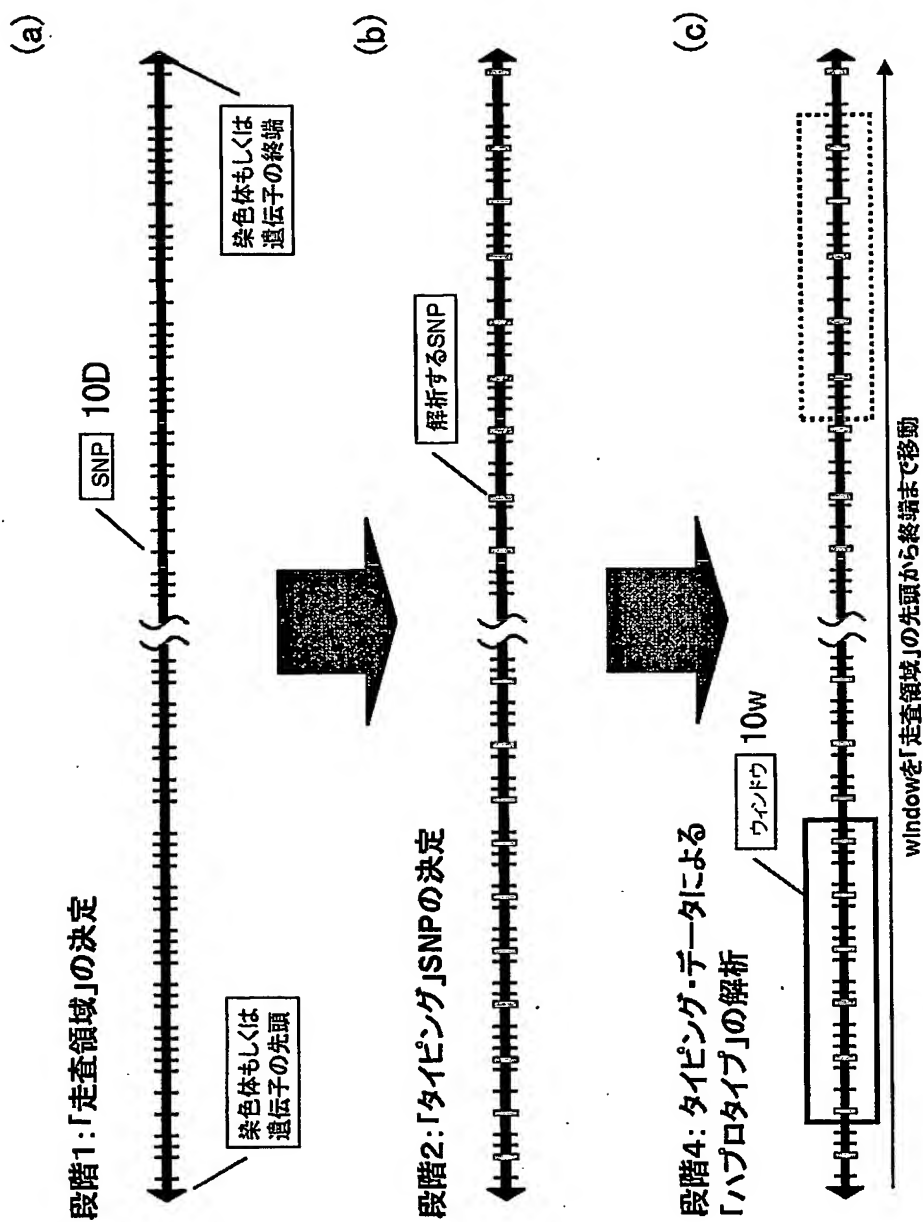
図面

【図 1】

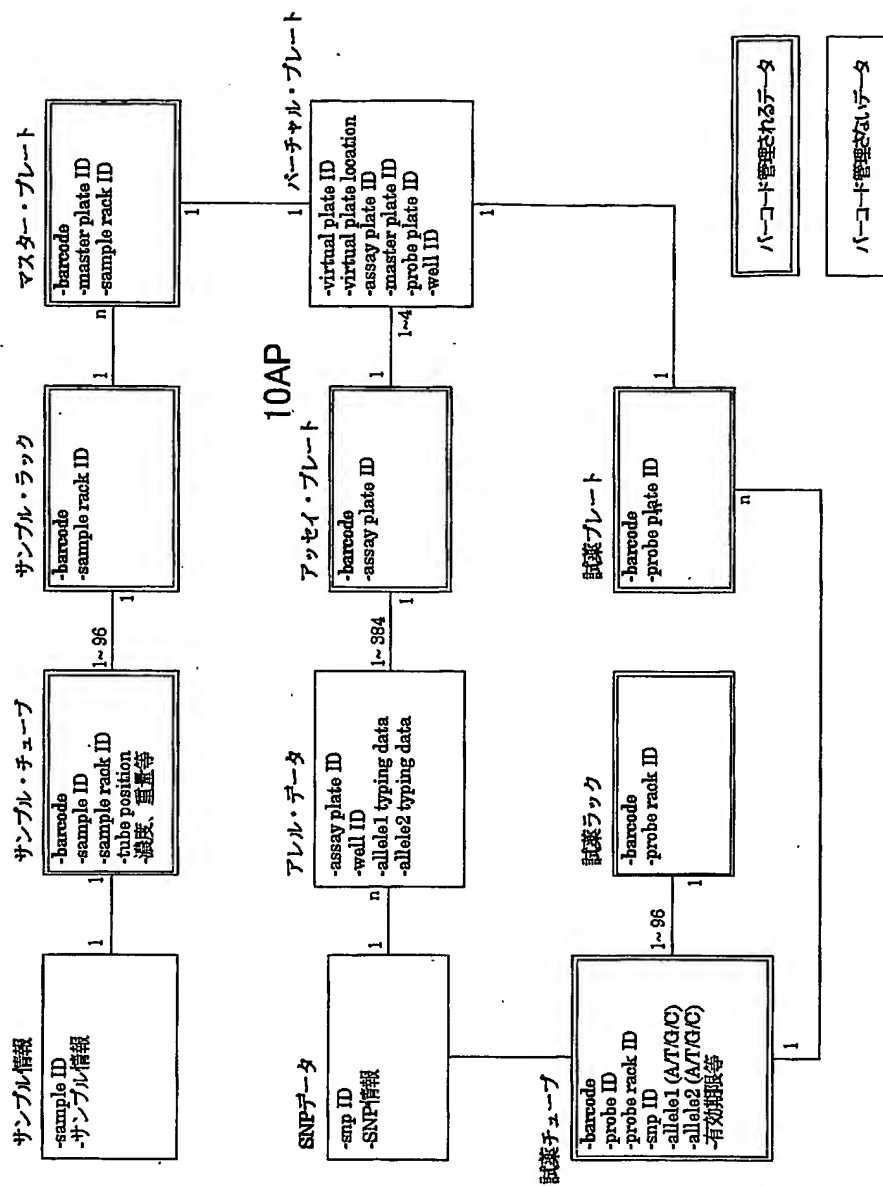
疾患易罹患性または薬剤応答性に関連するSNP特定のフロー



【図2】

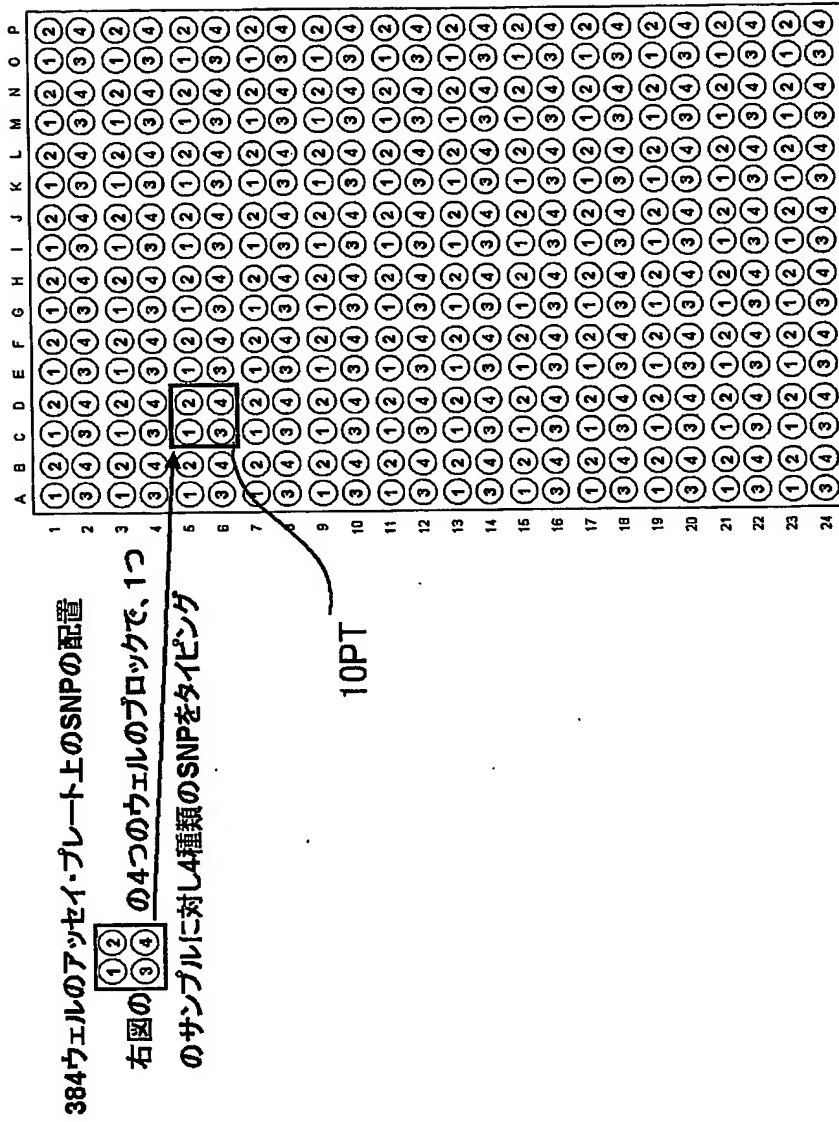


【圖 3】

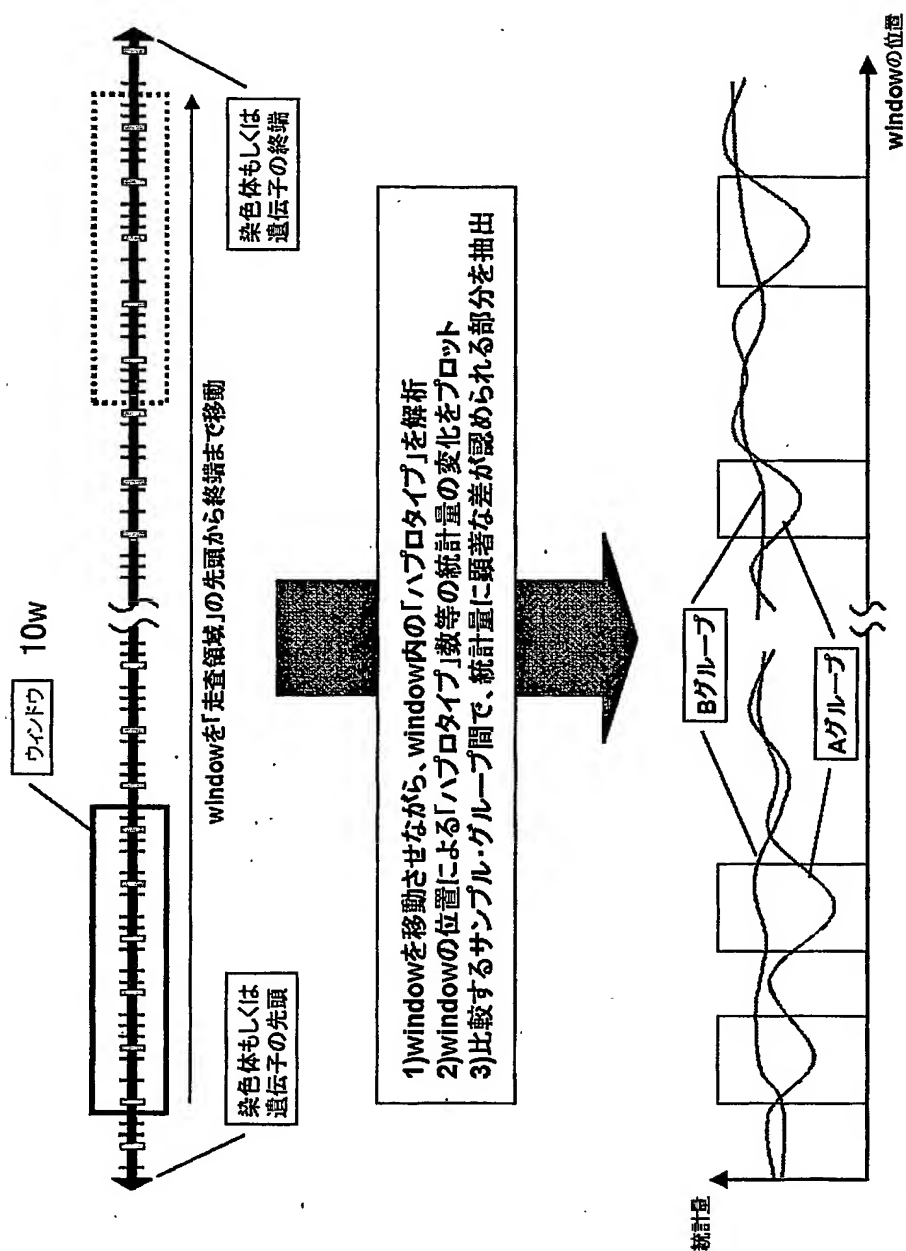




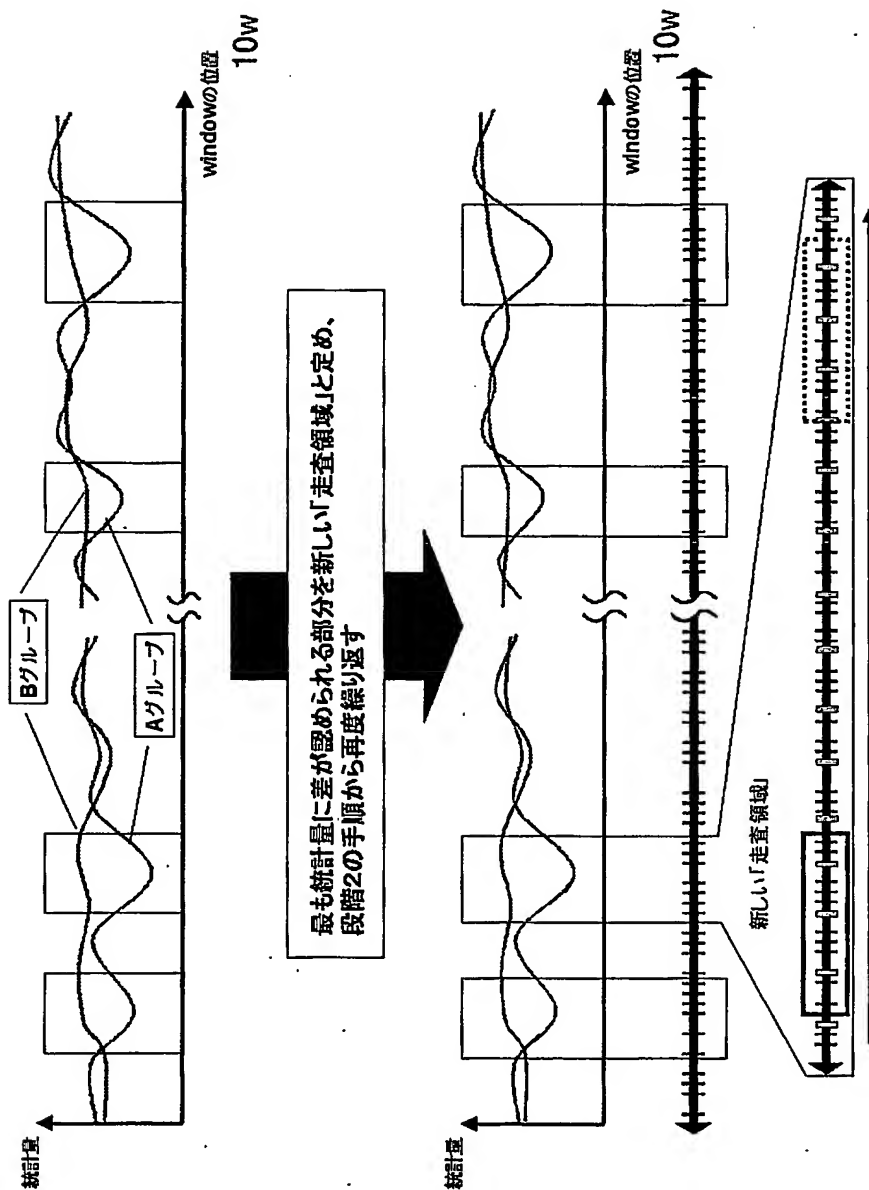
【図 4】



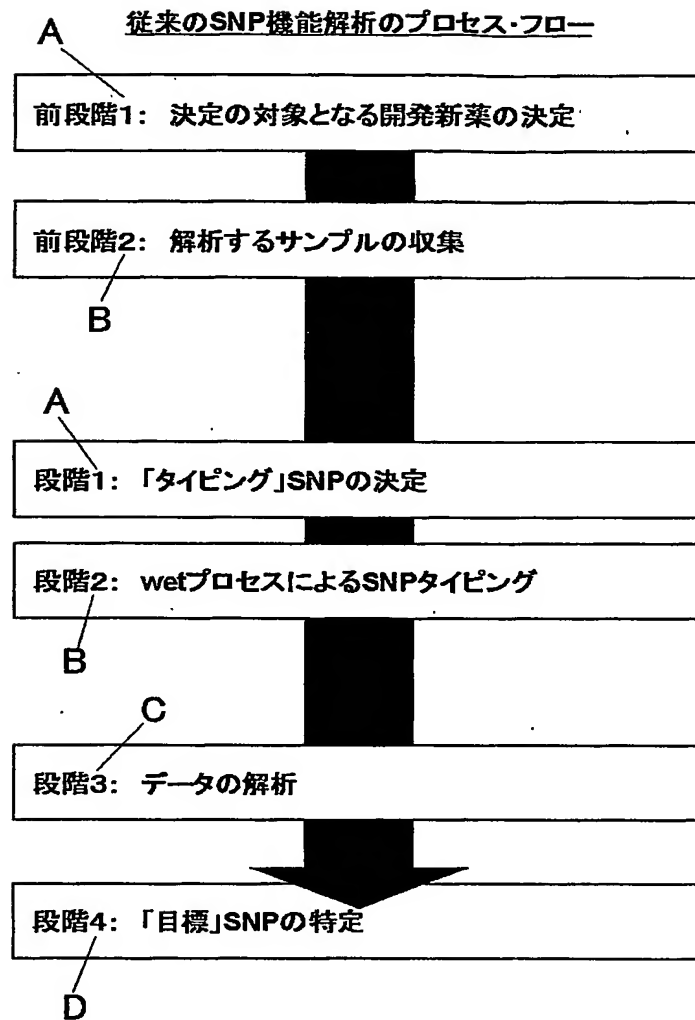
【図 5】



【図 6】



【図7】



【書類名】 要約書

【要約】

【課題】 ①マーカーとなるSNPの推定とその近傍SNPの詳細なタイピングを繰り返すことで、「目標」SNPが存在すると思われる塩基配列領域を段階的に絞込、最終的に「目標」SNPを効率良く特定するSNP特定方法に関する技術を提供する点にある。

【解決手段】 図1に示すように、本実施の形態1に係るSNP特定方法は、①決定の対象となる開発新薬の決定と②解析するサンプルの収集と③「走査領域（塩基配列領域）」の決定と④「タイピング」SNPの決定と⑤wetプロセスによるSNPタイピングとタイピング・データによるハプロタイプの解析と⑥「マーカー」SNPの推定（解析データの決定）と⑦「目標」SNP（目標SNP）の特定との工程を有し、①～⑦の工程をを1つのサイクル（処理サイクル）として繰り返す。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [502060360]

1. 変更年月日	2002年 1月15日
[変更理由]	新規登録
住 所	神奈川県藤沢市川名849の10クリオ藤沢参番館506
氏 名	鈴木 俊明

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**